

ABSTRACT OF THE DISCLOSURE

An object of the present invention is to suppress degradation of the quality in speech synthesis by selecting synthesis units so as to minimize a distortion caused by concatenation distortions and modification distortions. For that purpose, speech synthesis is performed by extracting a plurality of synthesis units corresponding to a phoneme environment from a synthesis-unit holding unit for holding a plurality of synthesis units so as to correspond to a predetermined prosody environment, calculating a distortion of each of the plurality of extracted synthesis units, obtaining a minimum distortion within a predetermined interval determined based on the prosody environment, selecting a series of synthesis units providing a minimum-distortion path, and modifying and concatenating the synthesis units.